

SEQ ID NO: 19

RESULT 7
G28289/c
LOCUS G28289 383 bp DNA STS 29-JUN-1996
DEFINITION human STS SHGC-34439, sequence tagged site.
ACCESSION G28289
VERSION G28289.1 GI:1397008
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 383)
AUTHORS Myers, R.M.
JOURNAL Unpublished (1996)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: CCACTGGCTAGAAGCCCTG
Primer B: TACCCACTGACTTCCAATCA
STS size: 126
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/uL
Total Vol: 10 uL

Buffer:
MgCl₂: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H4114
-- Washington University/Merck EST sequence.
FEATURES Location/Qualifiers
source 1. .383
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="12"
STS 21. .146
primer_bind 21. .39
primer_bind complement(125. .146)
BASE COUNT 101 a 73 c 95 g 107 t 7 others
ORIGIN

```

Qy 308 actgtggaatgaggatattttgaatcagttatggtaatgcacaaagcaatatttata 367
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |
Db 311 ACTGTGGAATGAGGATATTNTTGAATCAGTTATGGTAATGCAAAGCAATACTTATA 252

Qy 368 tgaacaacccaaagttagttctattnactgtactcttcgcccagtt 427
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |
Db 251 TGAANAAACCCAAAGTAGAGTTCTATTTAGCTGCTTATAACTGTACTCTCGCCCAGTT 192

Qy 428 caaaaaaaaagatttacatgacgtgccctgactgccaaagctccataccactgactctt 487
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |
Db 191 CAAAAAAAAGATTACATGACGTGCCCTGACTGCCNAAGCTCCATACCCACTGACTCTT 132

Qy 488 ccaatcaccaagtgtggaggctgccaccgagttcttgcaaataacaatgagaaca 547
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |
Db 131 CCAATCACCAAGTGCTGGAGGCTGCCACCGAGTCCTTGCAGAATACAACATGAGAAC 72

Qy 548 catccaaggcgtatttctcttcaagtca-ccagggttcttagccagtgggtggtcggc 606
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |
Db 71 CATCCAAGCAGTATTCTCTCTTCAAGTCACCCAGGGCTCTAGCCAGTGGGANNTNGGC 12

Qy 607 ctttcttactt 617
||| ||| ||| ||| |
Db 11 CCTTCTTACTT 1

```

```

RESULT 3
N39010
LOCUS N39010      542 bp   mRNA          EST      19-JAN-1996
DEFINITION yv22a08.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
IMAGE:243446 3', mRNA sequence.
ACCESSION N39010
VERSION N39010.1 GI:1162217
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 542)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 476
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: m13 -40 forward
High quality sequence stop: 476.
FEATURES Location/Qualifiers
source 1..542
/organism="Homo sapiens"
/db_xref="GDB:3792579"
/db_xref="taxon:9606"
/clone="IMAGE:243446"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
```

1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAAGAATTAAATTAAAGATCTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 135 a 159 c 123 g 124 t 1 others
ORIGIN

Query Match 42.4%; Score 486.4; DB 159; Length 542;
Best Local Similarity 99.6%; Pred. No. 1.1e-133;
Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 647 gtactaaatcccaggccagcagctgtcaacttcagtcccgactctgtgcctgtggc 706
Db 1 GTACTAAATCCCAGGCCAGCAGCTGTCACCTCAGTCCTCCGACTCTGTGCCTGTTGGTC 60

Qy 707 tttgcaaagggttctctgactcgaacacactggaaaaagttgtctctgtgactt 766
Db 61 TTTGCAAAGGTTCTCTGACTCGAACACACTGGAAAAGTTGTCTCTGTGACTTGTGACT 120

Qy 767 tcttgaaatcacaggctccagccactggaaagtggaaaactctgtgttaaccagaaaccta 826
Db 121 TCTTTGAATCACAGGCCAG-CACTGGAAAGTAAAAACTCTGCTGTTAACAGAAACCTA 179

Qy 827 caaacctcccaagggtggagaatcccagcagaaaaacacccccccaacagactccccct 886
Db 180 CAAACCTCCCAAGGTGGAAGAACATCCAGCAGAAAAACACCCCCCAACAGACTCCCCCT 239

Qy 887 ccaaagctggccaagaggatctgtccatatcttcctgactggatgataaaaattccc 946
Db 240 CCAAAGCTGGCCAAGAGGATCTGTCCAATATCTTCCTGACTGGATGATAAAAATTCCC 299

Qy 947 agaaaaaggccctcaggaggccttcctgtgcattggacctaaccacgaatccccagg 1006
Db 300 AGGAAAAGGGCCCTCAGGAGGCCTTCTGTGCATCTGGACCTAACGAAATCCCCAGG 359

Qy 1007 gagaaaccctggatatttccttccttcctggagcctatggaggagaagctgggtgtcc 1066
Db 360 GAGAAACCCCTGGATATTCCTTCCTTCCTGGAGCCTATGGAGGAGAAGCTGGTTGTCC 419

Qy 1067 tgcctttcccaaaagaaaaagcacgcactgctgagtgccaggccagccagaatgcc 1126
Db 420 TGCCTTTCCCCAGAGAAAAAGCACCGACTGCTGAGTGCCAGGGCCAGCCAGAACGCC 479

Qy 1127 gccctctgtccctccgcca 1146
Db 480 GCCCTCTGTCCCTCCGCCA 499

SEQ ID NO: 37

RESULT 10
G27766/c
LOCUS G27766 503 bp DNA STS 29-JUN-1996
DEFINITION human STS SHGC-33193, sequence tagged site.
ACCESSION G27766
VERSION G27766.1 GI:1396485
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 503)

AUTHORS Myers, R.M.
JOURNAL Unpublished (1996)
COMMENT

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: AGTATCTTATAAGTGACGTGGCTG

Primer B: GGGCTGCACTCTACCCTGTA

STS size: 125

PCR Profile:

Init

Denaturation: 94 degrees C for 15 seconds

Annealing: 62 degrees C for 23 seconds

Polymerization: 72 degrees C for 3

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Taq Polymerase: 0.05 units/ μ l
Total Vol.: 10 μ l

Total Vol: 10 ul

Buffer: MgCl₂ 2-5 mM

MgCl₂: 2.5 mM
KCl: 50 mM

Tris-HCl: 30 mM

pH: 8.3

pH: 8.3

Prepared with primer pairs provided by Sandoz

-- Washington University/Merck EST sequence.
Location/Qualifiers

	Location/Qualifiers
FEATURES	
source	1. .503 /organism="Homo sapiens" /db_xref="taxon:9606" /map="3"
STS	13. .137
primer_bind	13. .37
primer_bind	complement(118. .137)
BASE COUNT	118 a 119 c 145 g 112 t 9 others
ORIGIN	

Db	264	GCGATGGGACGATGGACAGAGACAGAGCGTGCANACGTAGAGTGGCTAGTGAAGGACGC	205
Qy	1297	cttttgcactttctggctcagcatgttgcattggaaataatgagactgagcc	1356
Db	204	CTTTTGACTCTCTGGCTCAGCATGTTGACTGGATTGAAATAATGAGACTGA-NC	146
Qy	1357	ctcggttggctgcacttacccgtacactgccttgaccctgagctgcaccc	1416
Db	145	CTCGGCTTGGGCTGCACCTCACCTGTANACTGCCTTGACCTGAGCTGCATCACCTCC	86
Qy	1417	ttaactgaggcgttcataaccatggagatgcctctttatgtcttcagccac-tcac	1474
Db	85	TAAACTGAGCAGTCTCATACCATGGAGAGATGCCTCTCTTATGTCTTCAGGCCACGTCA	26
Qy	1475	ttataaaagatacttatctttcagc	1499
Db	25	TTATAAAGATACTTCNCTTTCA	1

09445258Results

RESULT 5
A2HS_HUMAN
ID A2HS_HUMAN STANDARD; PRT; 367 AA.
AC P02765; O14962; O14961;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ALPHA-2-HS-GLYCOPROTEIN PRECURSOR (FETUIN-A) (ALPHA-2-Z-GLOBULIN)
DE (BA-ALPHA-2-GLYCOPROTEIN).
GN AHSG OR FETUA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87260816; PubMed=3474608;
RA Lee C.-C., Bowman B.H., Yang F.;
RT "Human alpha 2-HS-glycoprotein: the A and B chains with a connecting
RT sequence are encoded by a single mRNA transcript.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4403-4407(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97464058; PubMed=9322749;
RA Osawa M., Umetsu K., Sato M., Ohki T., Yukawa N., Suzuki T.,
RA Takeichi S.;
RT "Structure of the gene encoding human alpha 2-HS glycoprotein
RT (AHSG).";
RL Gene 196:121-125(1997).
RN [3]
RP SEQUENCE OF 19-300.
RX MEDLINE=86111834; PubMed=3944104;
RA Yoshioka Y., Gejyo F., Marti T., Rickli E.E., Burgi W., Offner G.D.,
RA Troxler R.F., Schmid K.;
RT "The complete amino acid sequence of the A-chain of human plasma
RT alpha 2HS-glycoprotein.";
RL J. Biol. Chem. 261:1665-1676(1986).
RN [4]
RP SEQUENCE OF 341-367.
RX MEDLINE=83161108; PubMed=6833285;
RA Gejyo F., Chang J.-L., Burgi W., Schmid K., Offner G.D., Troxler R.F.,
RA van Halbeek H., Dorland L., Gerwig G.J., Vliegenthart F.G.;
RT "Characterization of the B-chain of human plasma alpha
RT 2HS-glycoprotein. The complete amino acid sequence and primary
RT structure of its heteroglycan.";
RL J. Biol. Chem. 258:4966-4971(1983).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=89150282; PubMed=2645941;
RA Araki T., Yoshioka Y., Schmid K.;
RT "The position of the disulfide bonds in human plasma alpha 2 HS-
RT glycoprotein and the repeating double disulfide bonds in the domain
RT structure.";
RL Biochim. Biophys. Acta 994:195-199(1989).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=89340518; PubMed=2760061;
RA Kellerman J., Haupt H., Auerswald E.-A., Mueller-Esterl W.;
RT "The arrangement of disulfide loops in human alpha 2-HS glycoprotein.
RT Similarity to the disulfide bridge structures of cystatins and
RT kininogens.";
RL J. Biol. Chem. 264:14121-14128(1989).
RN [7]
RP SEQUENCE OF 34-367 FROM N.A., AND VARIANT AHSG*2.
RC TISSUE=Liver;
RX MEDLINE=97157105; PubMed=9003486;
RA Osawa M., Umetsu K., Ohki T., Nagasawa T., Suzuki T., Takeichi S.;
RT "Molecular evidence for human alpha 2-HS glycoprotein (AHSG)
RT polymorphism.";

RL Hum. Genet. 99:18-21(1997).
 CC -!- FUNCTION: PROMOTES ENDOCYTOSIS, POSSESSES OPSONIC PROPERTIES AND
 CC INFLUENCES THE MINERAL PHASE OF BONE. SHOWS AFFINITY FOR CALCIUM
 CC AND BARIUM IONS.
 CC -!- SUBUNIT: ALPHA-2-HS GLYCOPROTEIN DERIVES FROM THIS PRECURSOR,
 CC WHEN THE CONNECTING PEPTIDE IS CLEAVED OFF. THE TWO CHAINS A AND
 CC B ARE HELD TOGETHER BY A SINGLE DISULFIDE BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SELECTIVELY
 CC CONCENTRATED IN BONE MATRIX. IT IS ALSO FOUND IN DENTIN IN MUCH
 CC HIGHER QUANTITIES THAN OTHER PLASMA PROTEINS.
 CC -!- POLYMORPHISM: THERE ARE TWO COMMON ALLELES, AHSG*1 AND AHSG*2.
 CC AHSG*1 HAS THR-248 AND THR-256; AHSG*2 HAS MET-248 AND SER-256.
 CC -!- SIMILARITY: BELONGS TO THE FETUIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M16961; AAA51683.1; -.
 DR EMBL; D67013; BAA22652.1; -.
 DR EMBL; D67012; BAA22651.1; -.
 DR PIR; A29081; WOHU.
 DR PIR; S04467; S04467.
 DR PIR; S02765; S02765.
 DR GlycoSuiteDB; P02765; -.
 DR CarbBank; CCSD:11752; -.
 DR CarbBank; CCSD:36455; -.
 DR SWISS-2DPAGE; P02765; HUMAN.
 DR MIM; 138680; -.
 DR InterPro; IPR000010; -.
 DR InterPro; IPR001363; -.
 DR Pfam; PF00031; cystatin; 2.
 DR PROSITE; PS01254; FETUIN_1; 1.
 DR PROSITE; PS01255; FETUIN_2; 1.
 KW Glycoprotein; Signal; Repeat; Plasma; Mineral balance; Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 300 ALPHA-2-HS-GLYCOPROTEIN, CHAIN A.
 FT PEPTIDE 301 340 CONNECTING PEPTIDE.
 FT CHAIN 341 367 ALPHA-2-HS-GLYCOPROTEIN, CHAIN B.
 FT DOMAIN 27 144 CYSTATIN-LIKE 1.
 FT DOMAIN 145 260 CYSTATIN-LIKE 2.
 FT DISULFID 32 358 INTERCHAIN.
 FT DISULFID 89 100
 FT DISULFID 114 132
 FT DISULFID 146 149
 FT DISULFID 208 219
 FT DISULFID 230 247
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. .).
 FT CARBOHYD 256 256 O-LINKED (GALNAC. .).
 FT CARBOHYD 270 270 O-LINKED (GALNAC. .).
 FT CARBOHYD 346 346 O-LINKED (GALNAC. .).
 FT VARIANT 248 248 T -> M (IN AHSG*2).
 FT /FTId=VAR_002388.
 FT VARIANT 256 256 T -> S (IN AHSG*2).
 FT /FTId=VAR_002389.
 FT CONFLICT 16 16 C -> W (IN REF. 2).
 FT CONFLICT 54 54 W -> K (IN REF. 3).
 FT CONFLICT 125 125 F -> S (IN REF. 7).
 FT CONFLICT 204 204 S -> C (IN REF. 2).
 SQ SEQUENCE 367 AA; 39324 MW; 1AAF0C8D6B7E2789 CRC64;

Query Match 13.6%; Score 274; DB 1; Length 367;
 Best Local Similarity 26.2%; Pred. No. 4.2e-13;

	Matches	102;	Conservative	64;	Mismatches	169;	Indels	54;	Gaps	17;
Qy	11	ILVLCCGAMSPPQLALNPSALLSR--GCNDSDVLAAGFALRDINKDRKDGYVRLRNRVN	68							
	:	:: ::	:	: :	: :	: :	:	::		::
Db	5	VLLLCIAQLWGCHSAPHGPGLIYRQPNCDDPETEEAALVAIDYINQNLPWGYKHTLNQID	64							
Qy	69	DAQEYRRGGLGSLFYLTLDVLETDCHVLRKKAWQDCGMRIFFE-SVYGQCK-AIFYMNNP	126							
	:	:: :		:			:	:		::
Db	65	EVKVWPQQPSGELFEIEIDTLETTCHVLDPPTVARCSVRQLKEHAVEGDCDFQLLKLDGK	124							
Qy	127	SRVLYLAAYNCTLRPVSKKKIYMTCPDCPSSIPTDSSNHQVLEAAATESLAKYNNENTSKQ	186							
	:				:		:			:
Db	125	FSVVYA---KCDSSPDASAEDVRKVCQDCPLLAFLNDT--RVVHAAKAALAAFNAQNNGSN	179							
Qy	187	YSLFKVTRASSQWV-VGPSYFVEYLIKESPCT---KSQASSCSLQSSDSVPVGLCKGSLT	242							
	:		:: :		:		:	:		:
Db	180	FQLEEISRA--QLVPLPPPSTYVEFTVSGTDCVAKEATEAAKCNCNLLAEKQ--YGFCKATLS	235							
Qy	243	RTHWEKFVSVTCDFFESQAPATGSENSAVNQK-PTNLPKVEESQQKNTPPTDSPSKAGPR	301							
	:				:		:			
Db	236	EKLGGAEVAVTCTVFQTQPVTSQPPEGANEAVPT---PVVDPDAPPSPPLGAPG-LPPA	291							
Qy	302	GSVQYLPDLDDDKNSQEKGPQEAFPVHLDL-----TTNPQGETLDISFLFLEPMEEK	352							
	:			:						
Db	292	GSP---PDSHVLLAAPPGHQ-LHRAHYDLRHTFMGVVSLGSPSGE-----	332							
Qy	353	LVVLPFPKEKARTAACPGPAQNAPSPLVLP	381							
	:					:				
Db	333	--VSHP--KTRTVVQPSVGAAAGPVVPP	357							

RESULT : 6
A2HS_RABIT
ID A2HS_RABIT STANDARD; PRT; 360 AA.
AC P80191; O18997;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ALPHA-2-HS-GLYCOPROTEIN PRECURSOR (FETUIN-A) (HAEMONECTIN) (FRAGMENT).
GN AHSG OR FETUA.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;
RA Osawa M., Saito T., Takeichi S.;
RT "Nucleotide sequence of cDNA encoding rabbit fetuin.";
RL Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 19-26; 55-62; 118-128; 235-250 AND 332-344.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Bone marrow;
RX MEDLINE=93238730; PubMed=7682944;
RA White H., Totty N., Panayotou G.;
RT "Haemonectin, a granulocytic-cell-binding protein, is related to the plasma glycoprotein fetuin.";
RL Eur. J. Biochem. 213:523-528(1993).
CC -!- FUNCTION: A CELL ADHESION PROTEIN THAT BINDS IMMATURE CELLS OF THE GRANULOCYTE LINEAGE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: BONE MARROW.
CC -!- SIMILARITY: BELONGS TO THE FETUIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CYSTATIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D67014; BAA22653.1; -.
 DR PIR; S30340; S30340.
 DR InterPro; IPR000010; -.
 DR InterPro; IPR001363; -.
 DR Pfam; PF00031; cystatin; 2.
 DR PROSITE; PS01254; FETUIN_1; 1.
 DR PROSITE; PS01255; FETUIN_2; 1.
 KW Cell adhesion; Repeat; Glycoprotein; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 15 BY SIMILARITY.
 FT CHAIN 16 360 ALPHA-2-HS-GLYCOPROTEIN.
 FT DOMAIN 24 141 CYSTATIN-LIKE 1.
 FT DOMAIN 142 257 CYSTATIN-LIKE 2.
 FT DISULFID 29 351 BY SIMILARITY.
 FT DISULFID 86 97 BY SIMILARITY.
 FT DISULFID 111 129 BY SIMILARITY.
 FT DISULFID 143 146 BY SIMILARITY.
 FT DISULFID 205 216 BY SIMILARITY.
 FT DISULFID 227 244 BY SIMILARITY.
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 126 126 S -> P (IN REF. 1).
 SQ SEQUENCE 360 AA; 38386 MW; F0908740610CBB95 CRC64;

Query Match 13.4%; Score 270.5; DB 1; Length 360;
Best Local Similarity 25.0%; Pred. No. 7.4e-13;
Matches 99; Conservative 51; Mismatches 181; Indels 65; Gaps 13;

Qy	3	LLLPLALCILVLCCGAMSPPQLALNPSALLSRGCNDSDVLA VAGFALRDINKDRKDGYVL	62
: : : : : : : : : : :			
Db	1	LVLLLSLAQLWSCHLVTAVPLLGYR----EHNCDDPEAEQVAL LAUDHINNHLLQQGYKH	55
: : : : : : : : : : :			
Qy	63	RLNRVNDAQEYRRGGLGSLFYLTDVLETDCVLRKKAWQDCGMR- IFFESVYQGQCKAIF	121
: : : : : : : : : : :			
Db	56	ILNRIDKVKVWPRRPTGEVYELEIDTLETTCHALDPTPLANCSVR QVTQHAVEVGCD--F	113
: : : : : : : : : : :			
Qy	122	YMNPNPSRVLYLAAYNCTLRPVSKKKIYMTCPDCPSSIP TDSSNHQVLEAATESLAKYNNE	181
: : : : : : : : : : :			
Db	114	HVLKDQGQFTVLSAKCDSTPDSAEDILKLC PDCPLTLPLNNDT--RVAQAAEAALTA FNEK	171
: : : : : : : : : : :			
Qy	182	NTSKQYSLFKVTRASSQWVVG-P-SYFVEYLIKESP CTKSQA---SSCSLQSSDSVPVGLC	237
: : : : : : : : : : :			
Db	172	NNGAYLQLV EIARA--QLVPLPASTYVEFTVAATDCV AKEVTDPAKCNLLADKQ--YGFC	227
: : : : : : : : : : :			
Qy	238	KGSLTRTHWEKFVSVTCDF FESQ---APATGSE NSAVNQKPTNL PKVEESQQKNT PPTD	293
: : : : : : : : : : :			
Db	228	KATVAEKVAREEV EVTC TIFPAQP VVPQPQPG VAGAAA VEP---APAVD PASP VSPPDQG	284
: : : : : : : : : : :			
Qy	294	SPSK-----AGPRGSVQYLP DL DKNSQEKG PQEAFPV HLDLT NPQGETLDI	341
Db	285	SPSSLV VGPVL VAQAP APPRA HYDLR QTFAGV PMS EGSG EA F-----	327
Qy	342	SFLFLEPMEE KLVL LPFP KEKART AEC CPGPA QN ASP	377
: :			
Db	328	-----HPGK V P VV VQ P----- S VGA A P GP V IT PCP	352

RESULT : 1
US-08-737-045-14
; Sequence 14, Application US/08737045A
; Patent No. 5981483
; GENERAL INFORMATION:
; APPLICANT: Dennis, James W.
; APPLICANT: Denetriou, Michael
; APPLICANT: Mount Sinai Hospital Corporation

; TITLE OF INVENTION: COMPOSITIONS COMPRISING MODULATORS OF CYTOKINES OF THE
; TITLE OF INVENTION: TGF β SUPERFAMILY AND A METHOD OF TREATMENT WITH SUCH A
; TITLE OF INVENTION: COMPOSITION (AS AMENDED)
; FILE REFERENCE: 7933.94USWO
; CURRENT APPLICATION NUMBER: US/08/737,045A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-737-045-14

Query Match 13.6%; Score 274; DB 2; Length 367;
Best Local Similarity 26.2%; Pred. No. 1.1e-19;
Matches 102; Conservative 64; Mismatches 169; Indels 54; Gaps 17;

Qy 11 ILVLCCGAMSPPQLALNPSALLSR--GCNDSDVLAAGFALRDRINKDRKDGYVRLNRVN 68
Db 5 VLLLCLAQLWGCHSAPHGPGLIYRQPNCDDPETEEAALVAIDYINQNLPWGYKHTLNQID 64

Qy 69 DAQEYRRGGGLGSLFYLTLDVLETDCHVLRKKAWQDCGMRIFFE-SVYQQCK-AIFYMNNP 126
Db 65 EVKVWPQQPSGELFEIEIDTLETTCHVLDPTPVARCSVQLKEHAVEGDCDFQLLKLDGK 124

Qy 127 SRVLYLAAYNCLRPVSKKIYMTCPDCPSSIPTDSSNHQVLEAATESLAKYNNENTSKQ 186
Db 125 FSVVYA---KCDSSPDSAEDVRKVQCQDCPLLAPLNDT--RVVHAAKAALAAFNAQNNGSN 179

Qy 187 YSLFKVTRASSQWV-VGPSYFVEYLIKESPCT---KSQASSCSLQSSDSVPVGLCKGSLT 242
Db 180 FQLEEISRA--QLVPLPPSTYVEFTVSGTDCVAKEATEAAKCNLAAEKQ--YGFCKATLS 235

Qy 243 RTHWEKFVSVTCDFESQAPATGSENSAVNQK-PTNLPKVEESQQKNTPTDPSKAGPR 301
Db 236 EKLGGAEVAVTCTVFQTQPVTSQPQPEGANEAVPT---PVVDPDAPPSPPLGAPG-LPPA 291

Qy 302 GSVQYLPDLDDKNSQEKGPKQEAFPVHDL-----TTNPQGETLDISFLFLEPMEEK 352
Db 292 GSP---PDSHVLLAAPPGHQ-LHRAHYDLRHTFMGVVSLGPSGE----- 332

Qy 353 LVVLPFPKEKARTAECPGPAQNASPLVLP 381
Db 333 ---VSHPR-KTRTVVQPSVGAAAGPVVPP 357

RESULT 4
US-08-483-926A-10
; Sequence 10, Application US/08483926A
; Patent No. 5821227
; GENERAL INFORMATION:
; APPLICANT: Dennis, James W.
; TITLE OF INVENTION: MODULATORS OF CYTOKINES OF THE TGF BETA
; TITLE OF INVENTION: SUPERFAMILY AND METHODS FOR ASSAYING FOR SAME
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,926
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Sheep
US-08-483-926A-10

```

Query Match           13.2%; Score 265.5; DB 2; Length 364;
Best Local Similarity 29.5%; Pred. No. 7.7e-19;
Matches 92; Conservative 49; Mismatches 136; Indels 35; Gaps 15;

Qy      7 LALCILVLCCGAMSPPQLALNPSA-LLSRGCNDSDVLA VAGFALR DINKDRKDGYVRLN 65
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      6 LLFC LAQLC---SCRSIPLDPIAGYKEPACDDPDTEQA ALAAVDYINKHLPRGYKHTLN 61

Qy      66 RVNDAQEYRRGGLGSLFYLTLDVLETDCHVLRKKAWQDCGMRIFFE-SVYQCKA-IFYM 123
       :: : : | : | : : : | ||| | : | : | : | : | : | : | : | : | : | : |
Db      62 QIDSVKVWPRRPTGEVYDIEIDTLETTCHVLDPTPLVNCSVRQQTEHAVEVGDCDIHVLKQ 121

Qy      124 NNPSRVLYLAAYNCTLRPVSKKIYMTCPDCPSSIPTDSSNHQVLEAATESLAKYNNENT 183
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      122 DGQFSVLFT---KCDSSPD SAEDVRKLC PDCPLLAPL--NNSQV VHAAEVALATFNAQNN 176

Qy      184 SKQYSLFKVTRASSQWVVG-P SYFVEYLIKESPCTKSQA---SSCSLQSSDSVPVGLCKG 239
       : | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      177 GSYFQLVEISRA--QFVPLPGS VSVEFAVAATDCIAKEVVDPKCNLLAEKQ--YGFCKG 232

Qy      240 S-LTRTHWEKFVSUTCDF FES----QAPATGSENSAVNQKPTNLPKVEESQQKNTPTD 293
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      233 SVIQKALGGEDVTCTLFQTQPVIPQPQPEGAEAGA---PSAVPDAAVPD---AAVP 284

Qy      294 SPSKAG-PRGSV 304
       : | : | : | : |
Db      285 APSAAGL PVGSV 296

```

RESULT 15
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
; GENERAL INFORMATION:
; APPLICANT: Schmaier, Alvin H.
; APPLICANT: Jiang, Yongping
; TITLE OF INVENTION: Modulation of Blood
; TITLE OF INVENTION: Pressure and Inhibition of Platelet Activation
; TITLE OF INVENTION: with Kininogen Fragment
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna &
; ADDRESSEE: Monaco, P.C.
; STREET: 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,114B
; FILING DATE: 9 February 1994
; CLASSIFICATION: 514
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Application
; APPLICATION NUMBER: Serial No. 5472945 07/744,545
; FILING DATE: 13 August 1991
ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-137 CII
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5472945e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: peptide
; TOPOLOGY: linear
US-08-193-114B-1

Query Match 6.4%; Score 129.5; DB 1; Length 117;
Best Local Similarity 30.4%; Pred. No. 1e-05;
Matches 35; Conservative 18; Mismatches 43; Indels 19; Gaps 6;

Qy 151 CPDCPSSIPTDSSNHQVLEAATESLAKYNNENTSKQY-SLFKVTRASSQWVGPSYFVEY 209
| || ||||:| ::| | :| ||| :| :| :| | | | ||:::
Db 1 CVGCPRDIPTNSP--ELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDF 58

Qy 210 LIKESPCKSQ---ASSCSL---QSSDSVPVGLCKGSLTRTHWEK--FVSVC 254
: :|: |:| ||| ||| | :| ||| :| |
Db 59 VARETTCSKESNEELTESCETKKLGQLSD-----CNAEVYVVPWEKKIYPTVNC 107